

Latent Class Metric Conjoint Analysis

WAYNE S. DESARBO*

University of Michigan, Ann Arbor, MI 48109

MICHEL WEDEL

University of Groningen, P.O. Box 800, 9700 A.V. Groningen, The Netherlands

MARCO VRIENS

University of Groningen, P.O. Box 800, 9700 A.V. Groningen, The Netherlands

VENKATRAM RAMASWAMY

University of Michigan, School of Business, Ann Arbor, MI 48109

Key words: Conjoint Analysis, Mixtures of Distributions, Marketing Research, E-M Algorithm, Remote Entry Devices

[February, 1992]

Abstract

A latent class methodology for conjoint analysis is proposed, which simultaneously estimates market segment membership and part-worth utilities for each derived market segment using mixtures of multivariate conditional normal distributions. An E-M algorithm to estimate the parameters of these mixtures is briefly discussed. Finally, an application of the methodology to a commercial study (pretest) examining the design of a remote automobile entry device is presented.

In their reviews of the various applications of conjoint analysis, Cattin and Wittink (1982) and Wittink and Cattin (1989) have documented the increased commercial usage of conjoint analysis, as well as the increased number of research suppliers capable of performing such studies and the increased availability of computer software for customized study designs and analyses. The authors report that the majority of these applications involve the full profile method of stimulus construction with the use of metric rating scales, and a dramatic rise in the use of multiple regression for subsequent estimation (see Wittink and Cattin, 1989, table 1 on p. 92). Also mentioned was the fact that market segmentation ranks among the primary purposes for performing conjoint analysis in such commercial applications. Green and Krieger (1991) have also discussed how conjoint analysis can

*Wayne S. DeSarbo is the Sebastian S. Kresge Distinguished Professor of Marketing and Statistics at the School of Business of The University of Michigan in Ann Arbor, MI 48109. Michel Wedel is Professor of Marketing Research and Marco Vriens is Assistant Professor of Marketing in the Faculty of Economics at the University of Groningen in the Netherlands. Venkatram Ramaswamy is an Assistant Professor of Marketing at the School of Business of The University of Michigan in Ann Arbor, MI 48109.

be gainfully utilized for market segmentation purposes. Traditionally, this has been accomplished in one of two ways. In an a priori segmentation scheme (Green 1977), where segments are known or defined prior to the research, an aggregation of individual level preference judgments occurs and subsequent estimation of the conjoint model is performed at the segment level (cf., Green and Srinivasan 1978, 1990). In post hoc segmentation schemes (Green 1977), estimation of the conjoint analysis occurs at the individual level, and subsequent individual level part-worths are then clustered to form market segments. In the former scheme, problems typically arise since traditional demographic or psychographic background information that typically characterize market segments rarely adequately describe heterogeneous utility functions at the individual or market segment level (Moore 1980). In the latter scheme, the traditional two-stage methodology described will often influence the results obtained. In particular, the multiple regression and subsequent cluster analysis procedures typically optimize different and unrelated objective functions/aspects of the structure of the data. In addition, the classification literature (cf., Hartigan 1975) is abundant with evidence of how different clustering methods often produce different cluster (segment) results when applied to the same data. Finally, the conjoint analysis literature has well documented the potential instability of part-worth estimates derived at the individual level, especially in highly fractionated designs (cf., Wedel and Kistemaker 1989).

Recently, a number of procedures for performing market segmentation in conjoint analysis have been proposed. Hagerty (1985) developed a Q-factor analytic procedure that maximizes the predictive power of the derived segment level utility functions. This procedure models each consumer as, in part, belonging to every market segment and shows that the degree of membership of each consumer in each of the segments is determined by the first eigenvector of the correlation matrix calculated across the observed preference ratings among consumers. However, as noted by Kamakura (1988), such factor analytic procedures lead to overlapping clusters that are rarely identifiable. In addition, Stewart (1981) has argued that the number of factors obtained in such Q-factor analyses of individual characteristics (here, preferences) is not truly indicative of the true number of clusters. As Kamakura (1988) aptly notes, one may have more or less clusters than factors and the identification of the homogeneous clusters is subjective and complex, especially when there are more than two factors present. Finally, it must be mentioned that such factor analytic solutions identify prototypes and the correlations of individuals with these prototypes. These loadings are not equivalent to segment/prototype membership or probabilities of membership as they do not necessarily satisfy the row sum (to unity) constraints.

Kamakura (1988) developed a least-squares procedure for performing segmentation in conjoint analysis which attempts to group consumers into homogeneous segments so that their stated preferences are explained maximally by their group level preference functions. A two-step procedure is developed, conceptually similar to that of Spath (1985), where given a fixed number of segments and a binary indicator matrix designating membership of each individual in each market seg-

ment, part-worths are estimated in a least-squares fashion for each segment. Further, given these segment level part-worths, Kamakura (1988) develops an agglomerative hierarchical clustering procedure which attempts to optimize the same error sums-of-squares objective function. Two problems can be identified with this innovative approach. One, as pointed out by Kamakura (1988) himself, combining two individuals or clusters to form a new cluster early in this agglomerative scheme forces them to be in the same cluster in the latter stages of his algorithm. Thus, any misclassification in the earlier stages of the algorithm will be carried on to higher aggregation levels. Two, the hierarchical procedure imposes rather severe constraints on the aggregations: a hierarchy and the inability to allow overlapping or fuzzy cluster memberships. Note, Green and Helsen (1989) have shown that neither the Hagerty nor Kamakura approaches lead to higher predictive validity than are obtained by conventional conjoint analysis applied to individual response data.

Finally, a number of related approaches are also relevant to mention here. Ogawa (1987) has developed a stochastic, nonmetric approach for simultaneously estimating part-worths and aggregating individuals into segments. First, using a logit choice model framework, a ridgelike procedure for estimating individual part-worths is developed for *rank ordered* preferences. Then, an information-theoretic criterion or index is developed as a means to aggregate individuals. DeSarbo, Oliver, and Rangaswamy (1989) present a simulated annealing based methodology to perform clusterwise regression. These authors demonstrate the flexibility of their approach in accommodating multivariate measures, constraints on the resulting classification, and so on. Wedel and Kistemaker (1989) and Wedel and Steenkamp (1989, 1991) provide alternative clusterwise regression formulations for benefit segmentation that could also be adapted to perform such simultaneous estimation and segmentation in conjoint analysis.

Against this background, this paper presents an alternative latent class based specification for simultaneously performing conjoint analysis and market segmentation. The method employs a stochastic framework involving mixtures of multivariate conditional normal distributions allowing for heuristic tests for the optimal number of segments, fuzzy posterior probability of memberships that permit fractional membership in more than one market segment, and a stochastic approach (as opposed to deterministic, as in the vast majority of the other cases) that allows for asymptotic standard errors of the estimated part worths. In addition, the proposed methodology allows for heteroscedasticity among groups of consumers as well as for covariance within these groups' responses.

1. The latent class metric conjoint model

Let

$i = 1, \dots, I$ consumers;

$j = 1, \dots, J$ conjoint profiles;

- $l = 1, \dots, L$ independent variables in the conjoint analysis (coded typically as dummy variables);
- $k = 1, \dots, K$ derived segments;
- Δ_{ij} = the metric response rating to conjoint profile j elicited by consumer i ;
- X_{jl} = the value of the l -th independent conjoint variable in the j -th profile (for ease of exposition, we have dropped the distinction of factor and level within factor in this notation);
- β_{lk} = the estimated conjoint part-worth for the l -th independent conjoint variable for the k -th segment;
- $\underline{\Sigma}_k$ = a $J \times J$ covariance matrix estimated for segment k .

It is assumed that the row vector $\underline{\Delta}_i$ of dimension J has a probability density function which can be modeled as a finite mixture of the following conditional distributions:

$$H(\underline{\Delta}_i; \underline{\alpha}, \underline{X}, \underline{\beta}, \underline{\Sigma}) = \sum_{k=1}^K \alpha_k g_{ik}(\underline{\Delta}_i | \underline{X}, \underline{\beta}_k, \underline{\Sigma}_k), \tag{1}$$

where $\underline{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_{K-1})$ are the $K-1$ independent mixing proportions of the finite mixture such that:

$$0 \leq \alpha_k \leq 1, \text{ with } \alpha_k = 1 - \sum_{k=1}^{K-1} \alpha_k; \tag{2}$$

$\underline{X} = ((X_{jl})),$
 \underline{X}_j = the $1 \times L$ row vector of independent variables for the j -th profile;
 $\underline{\beta}_k$ = the $1 \times L$ row vector of conjoint part-worths for the k -th segment,
 $\underline{\beta} = ((\beta_{lk})),$
 $\underline{\Sigma}' = (\underline{\Sigma}_1, \underline{\Sigma}_2, \dots, \underline{\Sigma}_K).$

The distribution for each g_{ik} is specified as conditional multivariate normal:

$$g_{ik}(\underline{\Delta}_i | \underline{X}, \underline{\beta}_k, \underline{\Sigma}_k) = (2\pi)^{-J/2} |\underline{\Sigma}_k|^{-1/2} \exp[-1/2(\underline{\Delta}_i - \underline{X}\underline{\beta}'_k)\underline{\Sigma}_k^{-1}(\underline{\Delta}_i - \underline{X}\underline{\beta}'_k)']. \tag{3}$$

As seen, this framework is similar to the mixture framework in DeSarbo, Jedidi, Cool, and Schendel (1991) and DeSarbo, Howard, and Jedidi (1991) for simulta-

neously performing multidimensional scaling and cluster analysis, and generalizes DeSarbo and Cron (1988) to a multivariate setting. However, unlike the previous mixture/latent class approaches cited above that perform multidimensional scaling and a form of cluster analysis simultaneously and require nonlinear search methods at each iterate for parameter estimation (cf., DeSarbo, Jedidi, Cool, and Schendel, 1991), the present methodology performs a type of multivariate normal regression and a type of cluster analysis simultaneously, and analytical, closed-form, expressions can be derived for all parameter estimates within each iterate.

Given a sample of I independent consumers, one can thus form a likelihood expression:

$$L = \prod_{i=1}^I \left[\sum_{k=1}^K \alpha_k (2\pi)^{-J/2} |\underline{\Sigma}_k|^{-1/2} \exp\{-1/2(\underline{\Delta}_i - \underline{X}_i \underline{\beta}'_k) \underline{\Sigma}_k^{-1} (\underline{\Delta}_i - \underline{X}_i \underline{\beta}'_k)'\} \right] \tag{4}$$

or,

$$\ln L = \sum_{i=1}^I \ln \left[\sum_{k=1}^K \alpha_k g_{ik}(\underline{\Delta}_i | \underline{X}_i, \underline{\beta}_k, \underline{\Sigma}_k) \right]. \tag{5}$$

Given $\underline{\Delta}$, \underline{X} , and K , one wishes to estimate $\underline{\alpha}$, $\underline{\Sigma}$, and $\underline{\beta}$, in the full model so as to maximize expressions (4) or (5), given the conditions specified in equation (2).

1.1. Identification of group membership

Once estimates of $\underline{\alpha}$, $\underline{\Sigma}$, and $\underline{\beta}$ are obtained for any iteration of the maximum likelihood procedure, one can assign each consumer i to each latent class or market segment k via the estimated posterior probability (applying Bayes' rule), providing a fuzzy clustering:

$$\hat{P}_{ik} = \frac{\hat{\alpha}_k \hat{g}_{ik}(\underline{\Delta}_i | \underline{X}_i, \underline{\beta}_k, \underline{\Sigma}_k)}{\sum_{k=1}^K \hat{\alpha}_k \hat{g}_{ik}(\underline{\Delta}_i | \underline{X}_i, \underline{\beta}_k, \underline{\Sigma}_k)}, \text{ where } \sum_{k=1}^K \hat{P}_{ik} = 1, \text{ and } 0 \leq \hat{P}_{ik} \leq 1. \tag{6}$$

This particular feature allows for the flexibility of modeling situations where the members of a particular derived market segment display different preference patterns by allowing individuals to be fractional members of multiple segments. One could form partitions, if desired, by simply assigning consumer i to the market segment whose \hat{P}_{ik} was highest. Thus, our methodology will *simultaneously* estimate the conjoint part-worths and market segment membership. An Expectation-Maximization (E-M) algorithm (Dempster, Laird, and Rubin, 1977) is utilized for parameter estimation (see Ramaswamy, DeSarbo, Reibstein, and Robinson, 1992, who employ a similar mixture approach for pooling cross sections of time series observations with PIMS data in estimating marketing mix elasticities, for technical details).

1.2. Tests for T and K

As in DeSarbo, et al. (1991a, b), we use Akaike's (1974) information criterion (AIC) for the choice of the number of segments in such mixture clustering models. In our methodology, one would select K , which minimize:

$$AIC(K) = -2 \ln L + 2N(K), \quad (7)$$

where $N(K)$ is the number of independent parameters for the estimated model:

$$N(K) = KL + K[J(J-1)/2] + K - 1. \quad (8)$$

While this AIC test will be utilized to select appropriate values of K , this criterion relies essentially on the same regularity conditions needed for differences in $(-2 \ln L)$ to have its usual asymptotic distribution under the null-hypothesis (see Titterton, Smith and Makov, 1985). We therefore recommend the inspection of other goodness of fit measures such as a variance-accounted for (VAF) measure between $\hat{\Delta}$ and $\hat{\Delta} = \sum_{k=1}^K \hat{P}_{ik} \underline{X}_{ik}' \underline{\beta}'_k$ across all i , for values of K , as is done in traditional metric conjoint analysis. Also, one can construct an entropy-based measure:

$$E_K = 1 + \left[\sum_i \sum_k P_{ik} \ln P_{ik} \right] / I \ln K, \quad (9)$$

as suggested in Ramaswamy, et al. (1992), to examine centroid separation in the distributions. E_k is a relative measure bounded between 0 and 1. A value close to 0 indicates that the centroids of these conditional parametric distributions are not sufficiently separated for the number of segments specified.

2. Application

2.1. Study description

An automobile remote entry system allows a driver to lock or unlock a vehicle remotely within a certain distance of the vehicle by activating a small transmitter that is typically carried on a key chain. Several car manufacturers currently offer such remote entry system devices at an additional charge. A major automobile manufacturer was recently interested in determining which specific attributes to include in an automotive remote entry system from a consumer's perspective. Based upon discussions with consumers, product engineers, and marketing managers, seven attributes were included in this conjoint study:

1. *Type of Transmitter:* The first option is a multi-button transmitter that hangs from the key chain; each button activates a different feature of the system. The other option is a single-button on the vehicle ignition key that activates all the features of the system.
2. *Range of Operation:* The range of operation defines the maximum distance from the vehicle that the transmitter must be to make the system active. The options are either 10 feet or 30 feet.
3. *Feedback:* A remote entry system may or may not provide feedback to the operator. A system with feedback sounds the horn every time a button on the transmitter is depressed.
4. *Panic Alarm:* A remote entry system may or may not include a panic alarm feature. When activated, this feature sounds the horn and flashes the lights to indicate danger.
5. *Keypad:* A remote entry system may or may not include a keypad on the vehicle below the door handle. This keypad offers an alternate means of locking or unlocking the vehicle by punching into the keypad a unique five-digit code.
6. *Memory Features:* A remote entry system may or may not include memory features that automatically set the driver's seat and the power mirrors when the doors are unlocked with the use of the transmitter. Multiple transmitters for a given vehicle contain unique predefined settings for that driver.
7. *Trunk Release:* A remote entry system may or may not include a trunk release feature that unlocks the trunk when that feature is activated from the transmitter.

Table 1 presents the levels of the seven attributes with their respective codings. The table also provides the 2^7 fractional factorial design (Addelman, 1962) utilized for stimulus profile definition for main effects estimation. A pretest of this experiment was conducted with $N=48$ consumers. Pictures were provided to initially describe many of the features. In addition, to make the task more realistic, each feature level was assigned a cost or price, and each profile was appended with a

Table 1. Remote entry design attributes and design

	2 ⁷ Orthogonal Array							
	X-mit	Range	Feedback	Panic	Keypad	Memory	Trunk	Price*
Card 1	0	0	0	0	0	0	0	90
Card 2	0	0	0	0	1	1	0	240
Card 3	0	0	0	1	0	1	1	220
Card 4	0	0	0	1	1	0	1	170
Card 5	0	1	1	0	0	0	0	130
Card 6	0	1	1	0	1	1	0	280
Card 7	0	1	1	1	0	1	1	250
Card 8	0	1	1	1	1	0	1	200
Card 9	1	0	1	0	0	0	1	110
Card 10	1	0	1	0	1	1	1	260
Card 11	1	0	1	1	0	1	0	210
Card 12	1	0	1	1	1	0	0	160
Card 13	1	1	0	0	0	0	1	130
Card 14	1	1	0	0	1	1	1	280
Card 15	1	1	0	1	0	1	0	240
Card 16	1	1	0	1	1	0	0	190

*Total price was not included in the analyses; consumers only viewed total price and not individual attribute level costs/prices (shown below).

Attribute	Attribute levels
Transmitter	0 = Multi-Button (\$90) 1 = Single-Button (\$90)
Range	0 = 10 Feet 1 = 30 Feet (+ \$10)
Feedback	0 = No 1 = Yes (+ \$10)
Panic Alarm	0 = No 1 = Yes (+ \$20)
Keypad	0 = No 1 = Yes (+ \$50)
Memory Features	0 = No 1 = Yes (+ \$100)
Trunk Release	0 = No 1 = Yes (+ \$10)

total price that would be charged for such an entry system (to avoid the problem of the consumers uniformly preferring the profile containing the most features). Each consumer was shown the attribute levels and the total cost – not the individual attribute level costs. These total prices and attribute level costs are also displayed in table 1. Note, since total price is a linear combination of each of the

attributes, it was not included in the analysis explicitly. The consumer was asked to rate each profile using a 10-point preference scale.

2.2. Latent class metric conjoint results

Table 2 presents the statistical summary for K = 1 to 5 latent classes obtained with the latent class metric conjoint analysis methodology proposed earlier. Based on several computer runs, constraining Σ_k to be diagonal appears to represent the data most parsimoniously (according to the AIC statistic). According to the AIC (and VAF) criteria, K=4 classes appears to most parsimoniously describe the structure in these conjoint ratings. Table 3 presents the aggregate conjoint results over all N=48 consumers derived from the K = 1 solution (top right portion of the table). Here, the optimal design would include the multi-button transmitter, a 30-foot range, a panic alarm, no memory features, and a trunk release (cost = \$130). The feedback and keypad attributes are not significant overall. The overall VAF statistic is only 0.101. Table 3 also presents the estimated part-worths for the four class solution. Here, the first class finds the memory and trunk release features as most important, where preference is for no memory feature (perhaps due to its high cost) and a trunk release option. The second latent class finds the transmitter, range, panic alarm, memory, and trunk release features most important, where higher utility is derived for the single-button transmitter, 30 foot range, panic alarm, memory, and trunk release options at higher cost. The third latent class finds the panic alarm, key pad, memory, and trunk release options most preferred. Finally, the fourth latent class finds feedback, panic alarm, key pad, and memory all important, yet given that all the coefficients are negative, prefers a low cost remote entry device with few features. Thus, one observes rather dramatic differences between the aggregate part-worths and these four latent class results presented in table 3. In addition, the VAF statistic for the four class solution is over four times as large (0.414) as that of the K = 1 solution.

Table 3 also provides values of the estimated variances for each of the 16 profiles for each of the four latent classes. This table shows evidence of extreme

Table 2. Latent class metric conjoint results

K	Log-likelihood	DF	AIC	Entropy	VAF
1	-1764.21	24	3576.43	0.00	0.101
2	-1709.75	49	3517.49	0.94	0.257
3	-1685.55	74	3519.10	0.97	0.301
4	-1586.66	99	3371.32*	0.98	0.414
5	-1577.54	124	3403.09	0.90	0.463

*Denotes minimum AIC.

Table 3. Latent class metric conjoint K=4 results

	Latent Class				Aggregate
	k=1	k=2	k=3	k=4	
Intercept	5.73	2.83	1.76	9.28	4.14
Transmitter	0.35	-1.35**	-0.28	-0.22	-0.511**
Range	-0.08	0.91**	0.58	-0.32	0.41*
Feedback	0.07	-0.04	0.20	-1.54**	-0.10
Panic Alarm	0.14	1.46**	1.09**	-2.24**	0.62**
Keypad	0.33	0.38	0.79**	-3.28**	0.04
Memory	-2.40**	1.00**	-1.24**	-2.99**	-0.90**
Trunk Release	0.92**	1.26**	1.26**	-0.04	1.04**
α	0.31	0.38	0.19	0.12	1
n	15	18	9	6	48

*p<.05

**p<.01

Diag \sum

Profile	k=1	k=2	k=3	k=4
1	5.719	2.718	1.077	0.632
2	0.977	2.884	0.221	2.527
3	2.352	5.592	3.050	5.514
4	1.185	4.957	7.136	5.931
5	1.034	2.800	3.186	15.158
6	1.538	3.972	2.511	0.133
7	2.055	4.659	6.328	3.808
8	2.612	3.208	7.713	2.823
9	2.137	1.320	4.938	22.764
10	3.523	5.476	8.077	0.573
11	2.390	4.445	0.691	1.814
12	2.350	2.704	5.410	2.344
13	2.646	5.365	6.343	20.347
14	2.756	6.157	5.310	4.634
15	2.933	5.789	1.877	4.939
16	4.912	4.715	6.956	3.663

heteroscedasticity in the responses, especially in the fourth latent class concerning profiles 5, 9, and 13. The obvious inequality amongst these variance estimates both between and within latent classes attests to potential misspecification difficulties in applying other metric conjoint analysis techniques that assume homoscedasticity (the log likelihood drops to -2253.86 when assuming $\hat{\Sigma} = \sigma^2 \underline{I}$ for each latent class). The posterior probabilities for each of the N=48 consumers indicate good separation of the conditional mixture centroids as there is very little fuzziness ($E_K=0.98$) in the classifications. However, these probabilities of

class membership did not significantly relate to gender, previous ownership, or average yearly miles driven – the only three demographic variables initially collected.

2.3. Comparison with the traditional two-stage approach

Table 4 presents the individual level conjoint analysis coefficients obtained by multiple regression. As stated earlier, Wedel and Kistemaker (1989) and others have documented the potential instability of such individual part-worth estimates, especially in highly fractionated designs. Nonetheless, with the exception of consumers #1 and #48, and arguably #21 and #41, the main effects model appears to reasonably well approximate these consumers utility functions. Yet, as shown in table 4, there also appears to exist substantial heterogeneity in terms of the mixed signs of these estimated coefficients across the $N = 48$ consumers. The corresponding VAF statistic calculated over the entire sample is 0.902 (involving 768 parameters).

The traditional two-step approach entails first estimating such individual level coefficients (cf., table 4), and then clustering this sample of customers on the basis of these estimated individual level coefficients. However, as previously alluded to, a number of problems arise concerning this secondary clustering stage. Several decisions have to be made initially prior to obtaining clusters, and these decisions will affect the solutions obtained. For example, how should the data be preprocessed prior to analysis (if at all)? What metric should be selected to calculate proximity? Finally, what type of clustering should be used, and which amalgamation rule or algorithm should be employed? Most times there is no a priori theory to guide an analyst as to such decisions.

To illustrate some of these issues, we applied all seven clustering methods available in SYSTAT (Wilkinson, 1991): KMEANS, Ward's Method, Average Linkage, Complete Linkage, Centroid Method, Median Method, and Single Linkage, to the individual level part-worths estimated in table 4. To keep matters simple, we utilized the raw data (i.e., no preprocessing) and the Euclidean distance metric in the hierarchical clustering schemes. We then obtained four-cluster solutions for each of these seven methods and cross-tabulated the classifications with those derived from our latent class approach (after permuting label assignments to optimal congruence). Table 5 presents these cross-classifications and derived classification congruence percentages denoting the percent of the sample of $N = 48$ consumers that were assigned to the same clusters between our proposed latent class approach and each of the seven other methods (in pair-wise fashion). As seen in the table, these different clustering methods give rather dramatically different consumer classifications. Average linkage appears to best replicate the latent class assignments in allocating two-thirds of the sample to the same clusters. However, single linkage, the centroid method, and the median method solutions are plagued by chaining with many singleton clusters, accounting for less than

Table 4. Individual level estimated conjoint coefficients

Consumer	Intercept	Trans-mitter	Range	Feedback	Panic alarm	Keypad	Memory	Trunk release	R ²	Adj. R ²
1	1.250	0.000	0.000	-0.250	0.000	0.000	-0.250	0.250	0.429	-0.071
2	1.000	1.375	2.125	-1.125	-0.125	-0.125	-0.375	0.625	0.956	0.918
3	1.125	-0.125	1.125	-0.125	0.625	2.875	2.875	0.875	0.968	0.940
4	4.875	-0.250	-0.750	0.750	1.000	0.500	-1.750	1.500	0.821	0.665
5	5.125	0.250	-0.750	1.500	-2.000	-0.500	-2.000	2.250	0.952	0.910
6	2.750	-0.500	2.000	1.000	0.500	-0.750	-0.250	0.000	0.790	0.607
7	3.625	0.125	-0.125	-0.125	2.875	0.875	2.125	0.875	0.728	0.489
8	4.750	0.250	0.500	0.000	0.750	0.750	-5.250	0.750	0.971	0.946
9	5.500	-0.375	1.625	0.375	0.125	-0.125	-1.375	1.375	0.806	0.636
10	3.500	-1.000	2.000	-1.000	1.000	-1.500	1.500	1.500	0.965	0.934
11	2.000	-0.750	1.250	-1.000	1.250	1.000	-0.750	3.500	0.747	0.527
12	4.000	-2.000	1.000	0.250	0.000	1.000	0.250	-0.250	0.961	0.927
13	3.625	0.500	2.500	-0.250	-0.250	1.750	-2.750	1.500	0.978	0.959
14	8.625	0.500	-1.500	-0.750	-1.250	-2.250	-4.250	1.000	0.945	0.896
15	0.625	-1.625	0.125	0.125	2.125	0.875	3.125	2.125	0.899	0.810
16	6.500	-0.250	-0.250	-1.500	0.750	0.750	-3.500	0.500	0.955	0.916
17	1.250	0.625	2.625	-0.625	-0.375	2.125	0.875	0.875	0.847	0.713
18	3.000	0.125	-0.875	0.875	1.625	1.625	-0.125	1.375	0.804	0.633
19	5.750	0.625	0.625	-0.125	2.125	0.625	-5.625	0.375	0.968	0.939
20	1.375	2.250	2.500	-1.250	1.000	1.750	-1.750	1.750	0.961	0.927
21	3.625	1.625	0.625	0.375	0.375	0.375	-1.625	1.625	0.654	0.351
22	5.000	-0.750	0.000	0.250	2.250	0.000	1.250	0.500	0.909	0.830

Table 5. Cross classification frequencies

Latent Class (rows) By KMEANS (columns)						
	k = 1	k = 2	k = 3	k = 4	Total	Congruence Rate:
k = 1	11	4	0	0	15	58.3%
k = 2	0	15	3	0	18	
k = 3	1	7	1	0	9	
k = 4	5	0	0	1	6	
Total	17	26	4	1	48	
Latent Class (rows) By Ward (columns)						
	k = 1	k = 2	k = 3	k = 4	Total	Congruence Rate:
k = 1	7	0	2	6	15	60.4%
k = 2	0	10	8	0	18	
k = 3	2	1	6	0	9	
k = 4	0	0	0	6	6	
Total	9	11	16	12	48	
Latent Class (rows) By Average Linkage (columns)						
	k = 1	k = 2	k = 3	k = 4	Total	Congruence Rate:
k = 1	13	2	0	0	15	66.7%
k = 2	1	17	0	0	18	
k = 3	2	7	0	0	9	
k = 4	3	0	1	2	6	
Total	19	26	1	2	48	
Latent Class (rows) By Complete Linkage (columns)						
	k = 1	k = 2	k = 3	k = 4	Total	Congruence Rate:
k = 1	7	2	6	0	15	54.2%
k = 2	0	13	5	0	18	
k = 3	1	7	1	0	9	
k = 4	1	0	0	5	6	
Total	9	22	12	5	48	
Latent Class (rows) By Centroid (columns)						
	k = 1	k = 2	k = 3	k = 4	Total	Congruence Rate:
k = 1	0	15	0	0	15	43.8%
k = 2	0	17	1	0	18	
k = 3	0	9	0	0	9	
k = 4	1	1	0	4	6	
Total	1	42	1	4	48	
Latent Class (rows) By Median (columns)						
	k = 1	k = 2	k = 3	k = 4	Total	Congruence Rate:
k = 1	2	13	0	0	15	45.8%
k = 2	0	18	0	0	18	
k = 3	1	7	1	0	9	
k = 4	0	5	0	1	6	
Total	3	43	1	1	48	

Table 5 (continued).

	Latent Class (rows) By Single Linkage (columns)				Total	Congruence Rate:
	k = 1	k = 2	k = 3	k = 4		
k = 1	0	15	0	0	15	41.7%
k = 2	0	18	0	0	18	
k = 3	0	8	1	0	9	
k = 4	1	4	0	1	6	
Total	1	45	1	1	48	

50% common assignments with our latent-class results. Thus, the results from the proposed latent class conjoint analysis are substantially different from those obtained from the traditional two-stage conjoint analysis technique. In addition, fewer ad hoc decisions need to be made concerning analytical options as compared to these other approaches.

Finally, to further compare these diverse results, we took the means of the conjoint part-worths from the four-cluster solution derived from the Ward methodology (the second-best solution in terms of congruence with the latent class solution since the average linkage method produced small clusters with insufficient degrees of freedom for this comparison), and substituted them for β in the latent class procedure. Accordingly, we so redefined α and P with the assumption of homoscedasticity. The resulting \ln likelihood function was -2417.45 , much worse than the -1586.66 value obtained by the latent class solution. In addition, the corresponding VAF statistic for this Ward solution was 0.336, which is lower than that obtained with the $K=4$ latent class solution.

Acknowledgments

The authors are greatly indebted to James A. and Marcia T. Ciccarelli for their assistance on this project.

References

- Addelman, S. (1962). "Orthogonal Main-Effect Plans for Asymmetrical Factorial Experiments," *Technometrics* 4, 21-46.
- Akaike, H. (1974). "A New Look at Statistical Model Identification," *IEEE Transactions on Automatic Control* 6, 716-723.
- Bozdogan, H. (1987). "Model Selection and Akaike's Information Criterion: The General Theory and its Analytical Extensions," *Psychometrika* 52, 345-370.

- Cattin, Philippe, and Dick R. Wittink. (1982). "Commercial Use of Conjoint Analysis: A Survey," *Journal of Marketing* 44, 44–53.
- Dempster, A. P., N. M. Laird, and D. B. Rubin. (1977). "Maximum Likelihood from Incomplete data Via the E. M. Algorithm," *Journal of the Royal Statistical Society* B39, 1–38.
- DeSarbo, Wayne S., and William L. Cron. (1988). "Maximum Likelihood Methodology for Clusterwise Linear Regression," *Journal of Classification* 5, 249–282.
- DeSarbo, Wayne S., Richard L. Oliver, and Arvind Rangaswami. (1989). "A Simulated Annealing Methodology for Clusterwise Linear Regression," *Psychometrika* 54, 707–736.
- DeSarbo, Wayne S., Daniel J. Howard, and Kamel Jedidi. (1991a). "MULTICLUS: A New Method for Simultaneously Performing Multidimensional Scaling and Cluster Analysis," *Psychometrika* 56, 121–136.
- DeSarbo, Wayne S., Kamel Jedidi, Karel Cool, and Daniel Schendel. (1991b). "Simultaneous Multidimensional Unfolding and Clusters Analysis: An Investigation of Strategic Groups," *Marketing Letters* 2, 129–146.
- Green, Paul E. (1977). "A New Approach to Market Segmentation," *Business Horizons* 20, 61–73.
- Green, Paul E. (1984). "Hybrid Models for Conjoint Analysis: An Expository Review," *Journal of Marketing Research* 11, 155–169.
- Green, Paul E., and Kristiaan Helsen. (1989). "Cross-Validation Assessment of Alternatives to Individual-level Conjoint Analysis: A Case Study," *Journal of Marketing Research* 26, 346–350.
- Green, Paul E., and Abba M. Krieger. (1991). "Segmenting Markets with Conjoint Analysis," *Journal of Marketing* 55, 20–31.
- Green, Paul E., and V. Srinivasan. (1978). "Conjoint Analysis in Consumer Research: Issues and Outlook," *Journal of Consumer Research* 5 (September), 103–123.
- Green, Paul E., and V. Srinivasan. (1990). "Conjoint Analysis in Marketing Research: New Developments and Directions," *Journal of Marketing* 54, 3–19.
- Hagerty, Michael R. (1985). "Improving the Predictive Power of Conjoint Analysis: the Use of Factor Analysis and Cluster Analysis," *Journal of Marketing Research* 22 (May), 168–184.
- Hartigan, J. (1975). *Clustering Algorithms*. New York: Wiley.
- Kamakura, Wagner A. (1988). "A Least Squares Procedure for Benefit Segmentation with Conjoint Experiments," *Journal of Marketing Research* 25, 157–167.
- Moore, William L. (1980). "Levels of Aggregation in Conjoint Analysis: An Empirical Comparison," *Journal of Marketing Research* 17, 516–523.
- Ogawa, Kohsuke. (1987). "An Approach to Simultaneous Estimation and Segmentation in Conjoint Analysis," *Marketing Science* 6, 66–81.
- Ramaswamy, Venkatram, Wayne S. DeSarbo, David J. Reibstein, and William T. Robinson. (1991c). "A Latent Pooling Methodology for Regression Analysis with Limited Time Series of Cross Sections: A PIMS Data Application," Under Revision for *Marketing Science*.
- Spath, Helmut. (1985). *Cluster Dissection and Analysis*. New York: Springer-Verlag.
- Stewart, David W. (1981). "The Application and Misapplication of Factor Analysis in Marketing Research," *Journal of Marketing Research* 17, 51–62.
- Titterton, D. M., A. F. Smith, and V. E. Makov. (1985). *Statistical Analysis of Finite Mixture Distributions*. New York: Wiley.
- Wedel, Michel, and Cor Kistemaker. (1989). "Consumer Benefit Segmentation using Clusterwise Linear Regression," *International Journal of Research in Marketing* 6, 45–59.
- Wedel, Michel, and Jan-Benedict E.M. Steenkamp. (1989). "Fuzzy Clusterwise Regression Approach to Benefit Segmentation," *International Journal of Research in Marketing* 6, 241–258.
- Wedel, Michel, and Jan-Benedict E.M. Steenkamp. (1991). "A Clusterwise Regression Method for Simultaneous Fuzzy Market Structuring and Benefit Segmentation," *Journal of Marketing Research* 28, 385–396.
- Wilkinson, Leland. (1991). *SYSTAT: The System for Statistics*. Evanston, IL: SYSTAT, Inc.
- Wittink, Dick R., and Philippe Cattin. (1989). "Commercial Use of Conjoint Analysis: An Update," *Journal of Marketing* 53, 91–96.